We note our responses are indented and in blue.

> We thank the reviewer for these in-depth comments, they raise important points and are an important help in improving the article.

This work presents a technically ambitious and novel coupling of a three-dimensional functional–structural plant model (CPlantBox) with a microbial-explicit soil model (TraiRhizo) implemented in DuMux. The work addresses an important frontier—dynamically linking plant architecture, rhizosphere processes and microbial functioning to jointly simulate carbon and water fluxes across scales—and the overall concept and coding effort are impressive. The authors produce a rich set of scenario simulations that suggest compelling interactions between the timing of dry spells and microbial trait syndromes that can drive either net carbon retention or accelerated mineralization in the short term. The coupling and the multi-scale analysis framework are valuable contributions for modelling communities interested in rhizosphere processes and plant–soil feedbacks. To be honest, at 76 pages this is the most extensive review I've undertaken — it required a careful, time-consuming read to assess the modelling choices and their implications.

While the modelling effort is significant and presents useful, hypothesis-generating insights, the manuscript sometimes over-asserts ecological conclusions that the simulation scope and temporal extent cannot robustly support. Statements implying durable soil carbon "stabilisation" are particularly strong and risk overstating the results. The manuscript should consistently use language that reflects the model's temporal and mechanistic limits (e.g., "short-term net C balance shifts," "suggested potential for transient C accumulation," or "conditions that promote reduced short-term mineralization"), and explicitly frame the findings as model-derived hypotheses that require empirical validation over longer timeframes. Recasting conclusions in this way will preserve the impact of the study while avoiding claims that require organo-mineral interaction processes or multi-decadal validation that the current simulations do not include. My detailed comments and suggestions are described below.

> We will adapt the language and recast conclusions as recommended. We explain our intended adaptations in response to the detailed reviewer's comments below.

Abstract

Soil carbon stabilization is a long-term process involving complex organo-mineral interactions. Your model, as described, likely simulates microbial turnover and short-term respiration. To claim "stabilization" is a strong overstatement that must be tempered. The conclusions must be carefully rewritten to reflect exactly what the model outputs: namely, **net C flux** (input vs. respiratory output) and **short-term C storage** in microbial biomass or slow pools. The model may suggest a potential for stabilization, but it cannot conclusively prove it without validation against long-term experimental data.

> In the abstract and throughout the paper we will remove the word stabilization and use those two terms. We will also update the title of the article, following the recommendation of reviewer 1.
>
> In the abstract, we will update the sentences:
> - '[..] to include an implicit time-stepping.' To '[..] to include an implicit time-stepping. This allowed for an increase in the accuracy and stability of the model outputs.'

- ‘we observed a lasting stabilisation of soil C with less reactive communities.’ to ‘we observed an increase in the C storage of soil C with less reactive communities.’

Define what you mean by “earlier” versus “later” dry spells using absolute timing or plant phenological stages (e.g., days after sowing or leaf-stage) and by how much. Are these linked to specific phenological stages? What are the specific "soil biokinetic parametrisations"? Briefly state the difference between "reactive" and "less reactive" communities (e.g., differences in maximum growth rates or substrate use efficiencies).

We will add the start dates of the ‘earlier’ and ‘later’ dry spell and summarize in the abstract key parameters that differ between the three microbial parameter sets (growth rate, affinity to solutes, and yield).

Introduction

The introduction convincingly argues that coupled water–C dynamics at the soil–plant–atmosphere continuum are important under climate change, and it highlights the rhizosphere and microbes as key players. However, the justification for this specific modelling effort (why extend CPlantBox with TraiRhizo, why in DuMux, and why the chosen coupling is required now) is not tightly argued: the reader is told the model is useful, but not why existing tools cannot already answer the questions posed.

The introduction begins with very broad, textbook-level concepts (e.g., "water is a resource," "C is a building block") and takes too long to narrow its focus. The necessity of **your** work—integrating a **specific** microbial-explicit soil model (TraiRhizo) into a **specific** FSPM (CPlantBox) to address a **specific** gap—is buried and needs to be the central thread from the outset. The justification is currently implicit. For instance, the authors list several gaps (few models consider both water and C; many rhizosphere models ignore microbes; many FSPMs don’t simulate soil C), but they do not synthesize these into a single crisp gap statement that maps to the paper’s aim. The gap therefore reads as a set of related deficiencies rather than a targeted problem the paper will solve.

We will shorten the first section, to go quickly into the main limitations of current models and how our new coupling addresses them. The precise description of the soil and plant processes were meant to justify the processes implemented in the model, but this can be done more succinctly.

The text mentions updates (implicit time-stepping, updated multiscale coupling) but does not explain why these matter (stability? computational efficiency? ability to run larger/longer scenarios?). State concrete advantages and differences from prior CPlantBox couplings and from existing tools (e.g., TraiRhizo, SpatC).

We will update the text to underline the effect of these temporal scales: The implicit time-stepping and updated multiscale coupling lead to higher stability and accuracy in cases where we use CPlantBox to simulate plant and soil interactions.

The necessity of **your specific approach** (coupling CPlantBox with TraiRhizo) is not sufficiently justified. Why is this particular combination of models the best solution to the problem you've identified? A brief comparative sentence on the strengths of each model and the synergy of their coupling would be very effective.

We will add a direct explanation of how the two models can address current limitations listed in the introduction.

The two models were developed in parallel to respond to key gaps in existing models (simulations of soil and plant water and C, and of soil microbes) and to complement each other (similar levels of precision in the number of pools represented, complementary inputs and outputs).

Methods

The work described is highly sophisticated and represents a significant computational effort to integrate complex, multi-scale processes. The ambition to dynamically couple a 3D FSPM with a 1D microbial-explicit soil model is commendable.

Section 2.4.5: The definitions of the four different scales for analysis are innovative and thoughtful. The need to add a large, constant SOC_slow pool post-simulation because the model outputs were "almost always in the lowest SOC class" and "significantly below" measured values suggests a potential issue with the model's initial conditions or its fundamental ability to represent realistic background C levels. Adding SOCslow = SOCtheoric − SOCsimulated_init effectively injects an artificial, immobile C pool to match literature SOC. This is pragmatic, but it changes the meaning of “hotspot” thresholds and may bias hotspot fraction metrics. The assumptions and potential consequences must be clearly discussed and sensitivity tested. This requires a strong justification.

The effect of adding this pool will be discussed in more depth in the discussion section, as well as the limit of our calibration method. A sensitivity analysis and more precise calibration will be conducted for a following publication on the coupled model. The need to add this additional carbon pool was partly caused by the usage of the pre-set TraiRhizo parameter sets and process constraints.

Section 3.2 Why were the two one-week dry spells chosen at days 11–18 and 18–25 after sowing? Are these phenologically relevant for the virtual plant? Provide biological/ecological justification.

The selected days were chosen to have similar simulation scenarios to the one in Giraud et. al. (2023). These days were selected as they are within the vegetative stage of the plant: CPlantBox does not yet simulate production of flowers, fruits or seeds. The selected time spans correspond in our simulation to the seedling growth and tillering stages.

The phrase "warmer and drier atmospheric conditions were simulated" is too vague. Precisely which driving variables were changed (e.g., VPD, radiation, temperature) and by how much?

We will add a table providing changes in the environmental variables during the baseline conditions and the dry spells.

Results

The results presented are complex and stem from a highly sophisticated modelling effort. The multi-faceted analysis across different scales (plant, bulk soil, perirhizaltrunc, microscale) is a particular strength, providing a comprehensive view of the system's behaviour. However, the results are presented as absolute truths from the model without any acknowledgment of uncertainty or variability. Phrases like "we observed," "we found," and "led to" are used throughout, but the reader has no way of knowing: Are the differences between scenarios (e.g., earlyDry vs. lateDry) and parameter sets statistically significant or just numerically different? For instance, line 431-432 is is a key finding but is stated without any statistical test to back up the word "significant."

The connection between root architecture and C concentration peaks is well-described. However, the comparison between scenarios and parameter sets is again purely descriptive. A quantitative measure of the difference would be much more impactful.

Many statements are given as percentages or qualitative (“strong increase”, “higher”, “lowest”) but without absolute units or baseline magnitudes. The authors could replace qualitative descriptions with quantitative percentages, effect sizes, or other measures where possible.

> The reviewer raises a very good point. In the results section, we had wished to focus more on relative differences rather than absolute values, because of the limits of our calibration method, but it remains indeed necessary to offer a more precise description of the results. We will update this section, using values and statistical analysis to justify the causal effects described.

Strictly report what the model output is in the results section. Move all speculative explanations for why a pattern occurs (e.g., competition, rhizosphere overlap) to the Discussion section.

> We will move any interpretation of results to the discussion section.

Section 3.4, line 497: The claim that copiotrophs reach 580 µmol/cm³ is alarming high — show where and how large volumes have these concentrations (absolute volume in cm³) to contextualise (i.e., are hotspots tiny or spatially extensive?).

> A single truncated rhizosphere reached this value, which corresponds to 0.028651 cm^3. This high concentration is also linked to the fact that our model does not include the movement of microbes, a limitation mentioned in the discussion. We will underline the high values obtained in our discussion.

Discussion

The section 4.2 is overly long and reads more like an extended results section or a review of plant physiology. While comparing model outputs to literature is valid, the level of detail on biosynthetic growth, osmotic adjustment, and specific exudation rates from various papers is excessive and distracts from the paper's central focus on **soil-plant-microbe feedbacks**. This entire section should be drastically condensed. The key point—that the model qualitatively captured known dynamics despite its simplifications—can be made in a few paragraphs.

> The discussion of this section will be condensed (especially the comparison with the literature and the discussion of the different elements of growth). We will

discuss the changes compared with the results of Giraud et al (2023) and focus on the ‘key point’ of this section.

The introduction implicitly promised insights into how dry spell timing and microbial reactivity interact. While this is addressed in 4.3, the discussion does not explicitly return to frame the findings around the initial objectives. A strong discussion should begin by stating how the results have addressed the original knowledge gap. Additionally, Section 4.3 contains the most valuable discussion points but is buried. The insights about scale-dependency, microbial "starving-survival lifestyle," intra-microbial competition, and the rhizosphere priming effect are excellent and should form the core of the discussion.

More emphasis will be put on section 4.3, and parts of the current results section will be integrated. We will more directly reference the objectives given in the introduction, and use the new hypotheses added in the introduction section, following the recommendation of reviewer 1.

Several sentences imply general ecological truths (e.g., “the model showed… making a 3D evaluation relevant”), while the work actually shows model-based scenarios for a single soil/site/plant parameterisation set. Rephrase to emphasize that findings are model-derived insights that suggest hypotheses for empirical testing.

We will make our wording more careful throughout the discussion and conclusion sections.

The authors added an immobile SOC_slow pool to align simulated and literature SOC. This affects hotspot classification and interpretation. Discuss explicitly (a) how SOC_slow modifies hotspot thresholds and (b) whether the “stabilisation” conclusions hold without SOC_slow.

We will discuss the meaning and effect of the SOC_slow in more detail. Taking the advice from both reviewers 1 and 2 into account we shift the focus from carbon stabilisation to more general plant-soil interactions. Moreover, in the discussion part we will highlight the comparison between the scenarios, which makes the resulting bias caused by the SOC_slow less important for our conclusions.

Some paragraphs are long and mix methods/results/discussion. Keep the Discussion focused on interpretation, implications, limitations, and future work; move implementation details back to Methods or an Appendix.

We will reorganise the results and discussion section to make sure that there are no overlaps between methods, results and discussion.

Conclusion

The conclusion provides a brief and accurate summary of the work performed and correctly identifies the core findings. However, for a paper of this complexity and ambition, the conclusion is **significantly underdeveloped** and fails to adequately synthesize the study's full contributions, limitations, and broader implications. It reads more like an abstract than a conclusion. It should answer the "so what?" question. What is the broader significance of finding that C vs. water limitation in microbes dictates the soil C outcome? How does this advance the field of plant-soil modelling or inform future experimental work?

We thank the reviewer for this important and very helpful remark. We will replace the conclusion with the one given below, which aims to cover the points mentioned by Reviewer 2:

“In this paper, we presented the equations and implementation of a coupled model representing carbon and water flow in the soil-rhizosphere-plant continuum, influenced by atmospheric conditions through plant transpiration and photosynthesis. This framework accounts for the effects of water content variation on carbon flow and microbial activity. The multiscale implementation enables precise evaluation of fluxes and reactions at the soil-plant interface and can capture feedback across domains and scales. Despite the simplified representation of plant and microbial processes and limited calibration, the model reproduces trends reported in the literature, such as the “starving-survival” lifestyle of some microbial communities.

We found that the effect of water scarcity on soil carbon turnover is an emergent property arising from microscale feedbacks between plants, microbes, and local environmental conditions. The varied impacts of water stress highlight the strong influence of local soil and plant characteristics in determining whether carbon or water limitation dominates turnover. By representing multiple sources of plant and microbial stress, the model provides a mechanistic explanation for variability in plant and soil carbon allocation under combined stresses.

The model yields a wide range of variables (e.g., short-term carbon storage) across the soil-rhizosphere-plant continuum that can be used to assess the performance of plant phenotypes and management measures under dynamic conditions. By computing the resulting soil and plant conditions and identifying the key processes driving ecosystem responses, this model can both provide parameters for larger-scale models and inform the design of simplified models that focus on the most influential processes. Moreover, this work lays the foundation for a more comprehensive model of the plant-soil interaction cycle, including nutrient exchanges, which are essential to accurately represent the feedback of soil on plant processes. Finally, this model is uniquely adapted to re-create the experimental observations of isotopic carbon allocation in the plant and soil.”

The mention of experimental setup at the end of the conclusion is a reference to the work of Shultes et al. (2025, doi: 10.1038/s41467-025-62550-y), which will also be added in the discussion.

1. Typos

Fix small typos and duplicate words:

“to to the soil” -> “to the soil”;

“unterstanding” -> “understanding”;

“ans pore scales” -> “and pore scales”.

Line 399: “FSPM model of Giraud et al. (2023))” double parenthesis.

Line 433: respectivelly -> respectively.

Please check multiple typos and grammar issues:

“Expending” -> “Expanding”;

“compaired” -> “compared”;

“diurnal cylcing” -> “diurnal cycling”;

“hostpot” -> “hotspot”.

Thank you for taking the time to point out the spelling errors. These typos will be corrected and the text carefully checked for other errors.